# DESCRIPTIVE ANALYTICS OF AUTOMOTIVE SALES DATA

https://www.kaggle.com/datasets/ddosad/auto-sales-data

TABLE OF CONTENTS



## Reason for Selection of Data:

The Auto Sales dataset contains sales data of an Automobile company that sells automotive parts for different product categories including Classic and Vintage cars, trucks, ships and planes. It offers valuable insights into sales trends and pricing within the automotive industry. By analyzing this data, we can understand key business trends, like the demand for different product lines, which regions provide more sales, and pricing strategies, which are essential for decision-making in the automotive sector.

## Data Description:

The data contains the following columns:-

| Column Name | Description |
|---|---|
| ORDER NUMBER | This column represents the unique identification number assigned to each order. |
| QUANTITY ORDERED | It indicates the number of items ordered in each order. |
| PRICE EACH | This column specifies the price of each item in the order. |
| ORDER LINE NUMBER | It represents the line number of each item within an order. |
| SALES | This column denotes the total sales amount for each order, which is calculated by multiplying the quantity ordered by the price of each item. |
| ORDERDATE | It denotes the date on which the order was placed. |
| DAYS SINCE LAST ORDER | This column represents the number of days that have passed since the last order for each customer. It can be used to analyze customer purchasing patterns. |
| STATUS | It indicates the status of the order, such as "Shipped," "In Process," "Canceled”," "Disputed," "On Hold," or "Resolved." |
| PRODUCT LINE | This column specifies the product line categories to which each item belongs. |
| MSRP | It stands for Manufacturer's Suggested Retail Price and represents the suggested selling price for each item. |
| PRODUCT CODE | This column represents the unique code assigned to each product. |
| CUSTOMER NAME | It denotes the name of the customer who placed the order. |
| PHONE | This column contains the contact phone number for the customer. |
| ADDRESS LINE 1 | It represents the first line of the customer's address. |
| CITY | This column specifies the city where the customer is located. |
| POSTAL CODE | It denotes the postal code or ZIP code associated with the customer's address. |
| COUNTRY | This column indicates the country where the customer is located. |
| CONTACT LAST NAME | It represents the last name of the contact person associated with the customer. |

| | |
|---|---|
| CONTACT FIRST NAME | This column denotes the first name of the contact person associated with the customer. |
| DEALSIZE | It indicates the size of the deal or order, which are the categories "Small," "Medium," or "Large." |

## What do I intend to evaluate?

Below are a few areas to explore through descriptive analytics

1. Investigate relationship between Pricing and Order Volume-
   This allows us to understand how price affects the quantity ordered. This can help the automotive business industry to adjust prices to optimize sales.

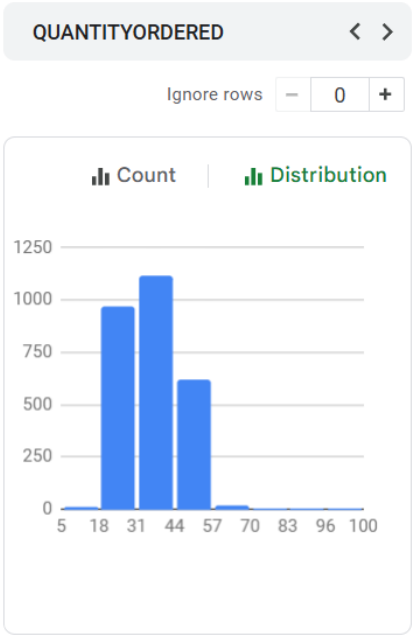
2. Country wise sales -
   Understanding trends in sales patterns across different cities is necessary to assess regional demand. This can help to target customers based on their demographics and offer effective pricing models according to the country they belong.

3. Analyze sales by product line-
   It is necessary to analyze sales patterns across the different product lines and identify which products contribute to more sales. This is done so as to recognize the products that are highly in demand and hence can focus on inventory and marketing strategies that will generate greater revenue for the business.

## Descriptive statistics

| | QUANTITY ORDERED | SALES | PRICEEACH | DAYS_SINCE_LASTORDER |
|---|---|---|---|---|
| Mean | 35.10302148 | 3553.047583 | 101.0989516 | 1757.085912 |
| Standard Error | 0.1862582415 | 35.08661833 | 0.8021576172 | 15.63159624 |
| Median | 35 | 3184.8 | 95.55 | 1761 |
| Mode | 34 | 8209.44 | 96.34 | 2207 |
| Standard Deviation | 9.762135424 | 1838.953901 | 42.04254925 | 819.2805763 |
| Sample Variance | 95.29928803 | 3381751.448 | 1767.575947 | 671220.6627 |
| Kurtosis | 0.4428649615 | 1.773100084 | 0.2285186185 | -1.02446586 |
| Skewness | 0.3692863511 | 1.155939788 | 0.6972217651 | -0.002983408147 |
| Range | 91 | 13600.67 | 225.99 | 3520 |
| Minimum | 6 | 482.13 | 26.88 | 42 |
| Maximum | 97 | 14082.8 | 252.87 | 3562 |
| Sum | 96428 | 9760221.71 | 277718.82 | 4826715 |
| Count | 2747 | 2747 | 2747 | 2747 |
| Largest(1) | 97 | 14082.8 | 252.87 | 3562 |
| Smallest(1) | 6 | 482.13 | 26.88 | 42 |
| Confidence Level(9 | 0.3652204138 | 68.79883094 | 1.572893283 | 30.65087485 |

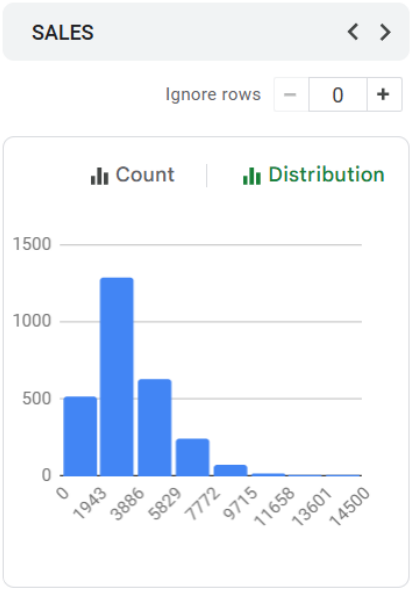
## 1.Quantity Ordered




- **Mean**: The average quantity ordered is around 35 units, which indicates that customers typically order products of around this quantity.
- **Standard Deviation**: With a standard deviation of 9.76, there is a moderate spread in the quantity ordered, indicating some variability in orders.

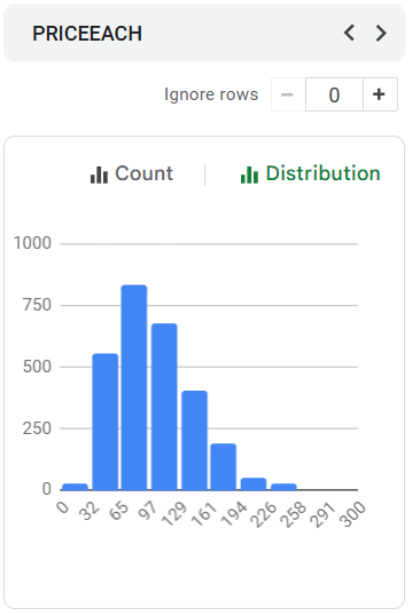
- **Range**: The range of 91 (from 6 to 97 units) indicates that there is a significant variation in orders, with some customers ordering very small quantities of parts and others ordering nearly 100 units of some parts.
- **Skewness & Kurtosis**: The skewness (0.37) and kurtosis (0.44) values show a right-skewed and nearly normal distribution. This tells us that there are a few high quantity orders but no extreme outliers.

## 2. Sales



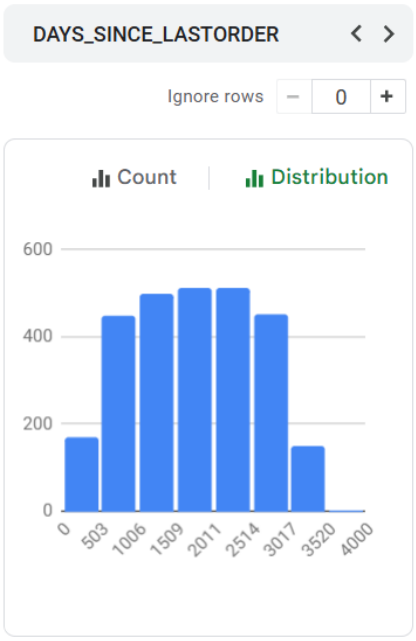

- **Mean**: The average sales amount per order is approximately 3,553. This is the average value around which most sales amounts are distributed.
- **Standard Deviation**: A standard deviation of 1,838.95 shows that there is a high variance in sales, which aligns with the large range of sales i.e. from 482.13 to 14,082. This spread indicates that some sales are much larger or smaller than others.
- **Range**: The high range of 13,600.67 shows us the varied nature of orders in terms of revenue generation, from low-revenue orders to high-revenue ones.
- **Skewness & Kurtosis**: A skewness of 1.15 and kurtosis of 1.77 indicate a positively skewed distribution, with a concentration of lower sales outliers and some higher-value outliers. This may indicate a few high-value orders, while most orders generate lower revenue.

## 3. Price Each




- **Mean**: The average price per item is around 101.1, providing a central benchmark for item pricing.
- **Standard Deviation**: With a standard deviation of 42.04, there is considerable variability in part prices, which could be due to a range of products for different product lines with different values.
- **Range**: The range of 225.99 (from 26.88 to 252.87) further reflects a wide range of product prices, from inexpensive to premium parts for the different product categories.
- **Skewness & Kurtosis**: The skewness of 0.70 suggests a moderate positive skew, indicating that while most items are priced close to the average, there are few higher-priced products as well. The low kurtosis (0.23) shows a fairly normal distribution without heavy tails.

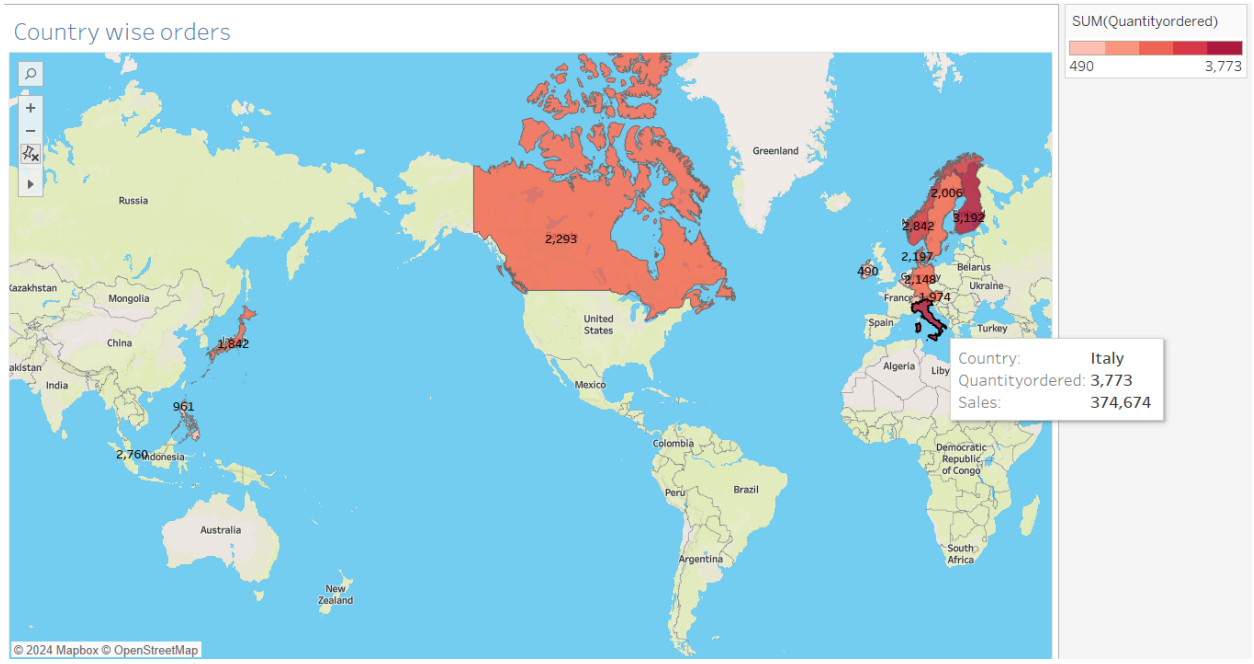
## 4. Days Since Last Order




- **Mean**: The mean days since the last order is 1,757, suggesting that there is considerable gap since customers' previous orders.
- **Standard Deviation**: The standard deviation of 819.28 days indicates high variability in how recently customers placed their last orders, with some customers ordering frequently and others very seldom.
- **Range**: The range of 3,520 days (from 42 to 3,562 days) shows that some customers have been inactive for many years, while others ordered recently. This wide range could indicate a mix of long-term customers and newer customers or even customers with changed interests.
- **Skewness & Kurtosis**: The negative kurtosis (-1.02) and almost zero skewness (-0.003) suggest a distribution that is close to normal but slightly flatter than a typical bell curve, with no extreme outliers.

## Visualizations

### 1.Country wise orders

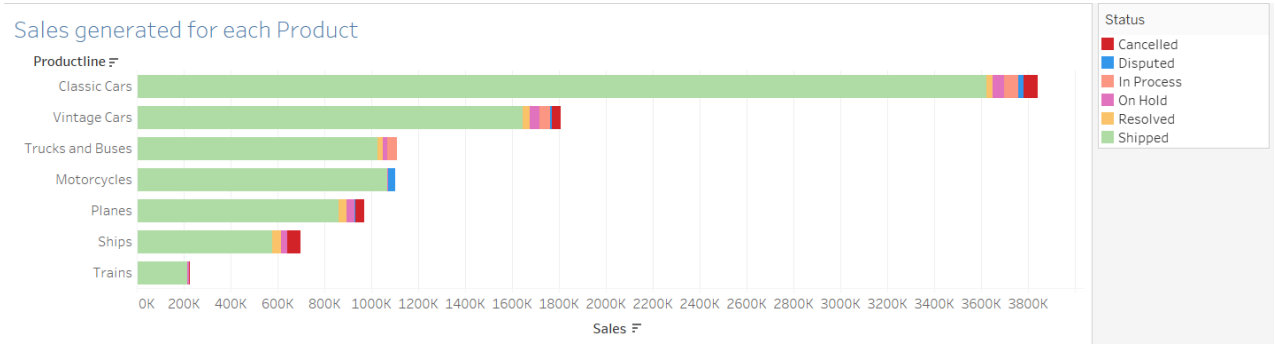




- This visual shows us region wise orders, we observe that Italy has ordered the highest amount of products i.e 3773 hence generating the highest sales for the company i.e a total of 374674
- Followed by Finland that has ordered 3192 products. These are our high valued customers and should be prioritized for retention efforts to maintain their loyalty.

- On the other hand Ireland has ordered just 490 products and Philippines has ordered 961 products, hence efficient marketing strategies need to be deployed in these regions in order to help attract more customers and drive sales.
- It is also observed that many parts of Europe have purchased products from the company apart from some cities in the USA and Australia , Singapore and Japan in the Eastern hemisphere. This perhaps showcases adequate marketing efforts and demand for these products in Europe compared to other parts of the world.
- The visual below shows region wise sales, it is observed that among countries the United States leads in sales, followed by Spain and France. Among cities, Madrid tops sales with 1082k, followed by San Diego and New York City.

## 2. Sales generated for each Product Line




This visualization shows the sales generated for each product category, along with the status of orders in each category.

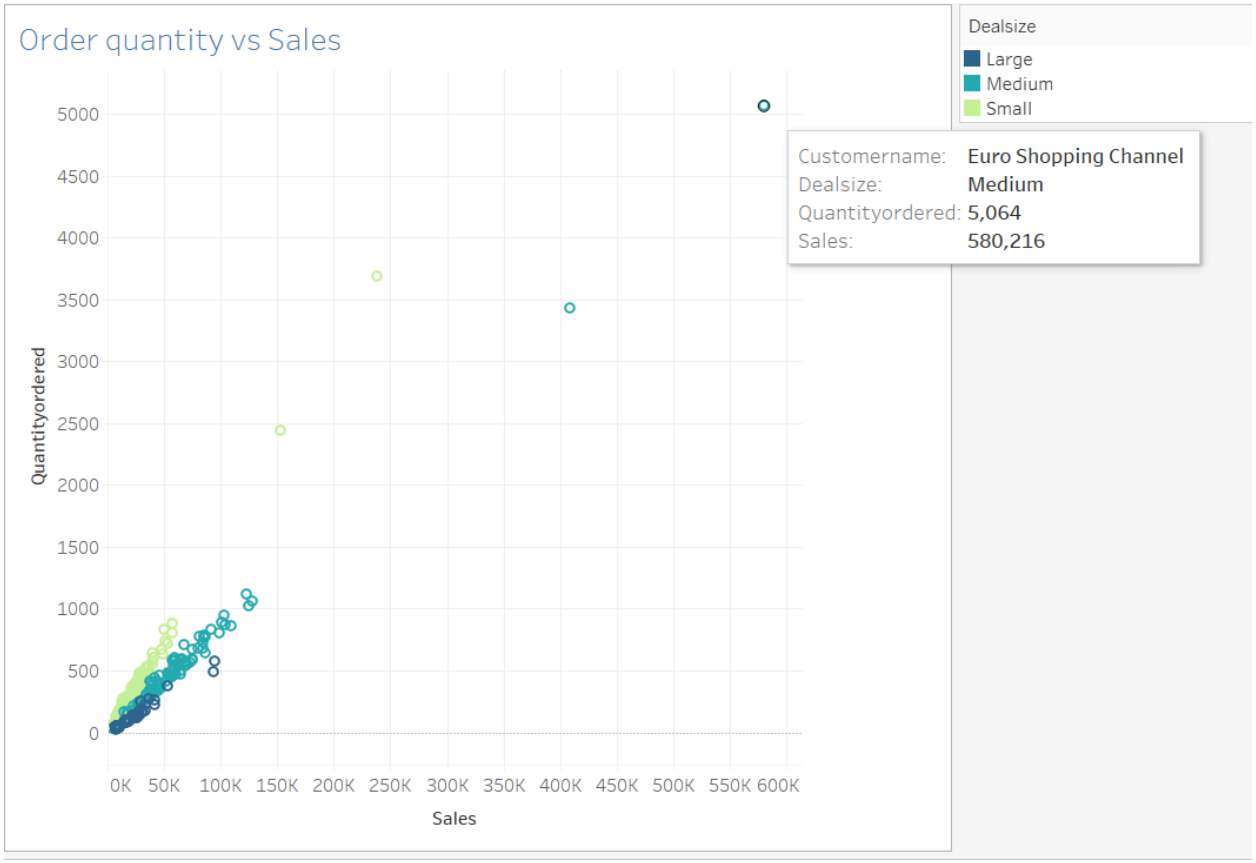
- It is observed that Classic Cars generate the highest sales by a huge amount i.e 3,842K . Most of these products have been shipped, indicating that customers are fully satisfied. Vintage Cars also perform well, with a sales of 1,806,676 , here too most of the products have been shipped.
- Trucks, Buses, motorcycles, planes and ships generate moderate sales, they do contribute to the revenue significantly with sales close to 1000K each. Planes and ships have a slightly higher occurrence of On Hold and Canceled statuses, which might be due to some issues.
- Trains generate the least amount of sales , with revenue around 200k. This category has a high proportion of “Canceled” orders, suggesting changes in demand.
- Overall, most of the products have been shipped, showing a strong overall fulfillment rate. However, Canceled and On Hold statuses prevalent in Ships, planes and trains could be due to factors like production delays, supply chain issues, lower customer interest or changes in demand.

## 3. Scatter/ Box Plot of MSRP for each product


This box plot visualizes the Manufacturer's Suggested Retail Price (MSRP) for different product lines, segmented by deal size. Each product line has its own spread of MSRP values.

- In most cases the Price Each is nearly equal to the MSRP, while in lower MSRP categories like Trains, the actual price is lower than the MSRP. This may suggest discounts or price adjustments for this product line.
- Understanding of Outliers: Classic Cars have the widest range of MSRP values, with several outliers above 2500, indicating premium-priced products in this category.
- Motorcycles, Planes and Ships have MSRP values more tightly clustered around lower values below 1500. Trains show low MSRP values, with fewer outliers in higher price ranges.
- It is seen that Mini Gifts Distributors Ltd. purchased a Classic Car product with a medium deal size, priced at a high MSRP.

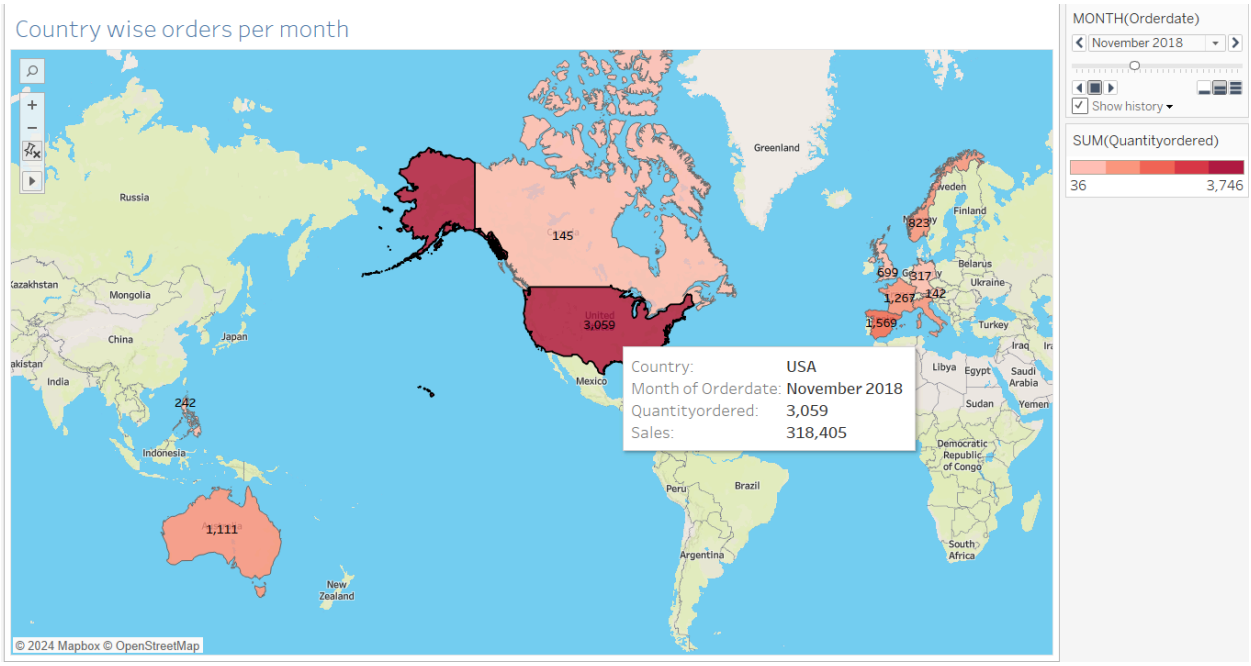
## 4.Scatter plot of Quantity Ordered and Sales




- When customers order more items (QUANTITY ORDERED), it's often associated with higher sales (SALES). So, the more customers buy, the more the company earns in sales.
- Here we see the relationship between the amount ordered and sales. It is observed that one customer 'Euro shopping Channel' ordered a high quantity and hence led to a high revenue for the company.

# Application of pages and filters

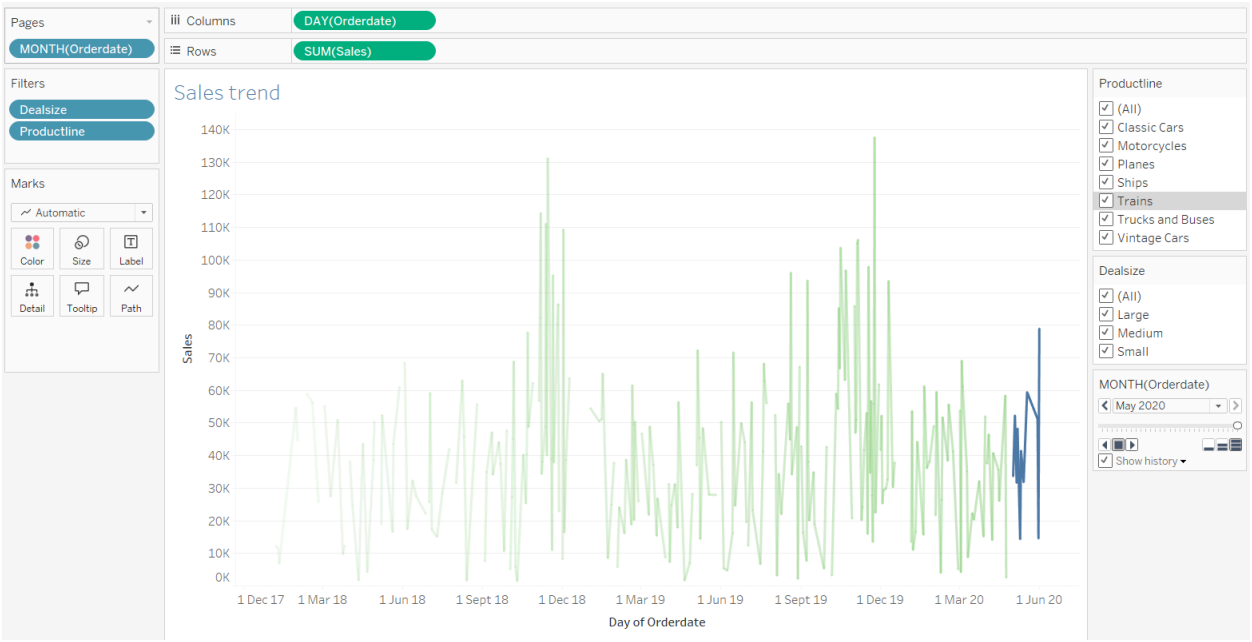
1. Country wise orders per month




**Country wise quantity trend analysis with Pages:** This is a visualization of Country wise orders per month, it shows monthly how the orders change using the Pages functionality of Tableau.

- It is observed that in November 2018 the quantity ordered in the USA was very high 3059 compared to the other months. Similar trend was observed in 2019 November wherein the quantity ordered was 3746.
- Similarly seasonal trends are observed monthly every year, the quantity ordered follows similar patterns in the same months.
- However, there is a surge in quantity ordered in the 1st quarter of 2020 which is similar to the seasonal orders in the month of November, this could be due to the start of the COVID-19 pandemic in 2020 which led to changes in consumer habits which may have contributed to higher sales activity during this period.

## 2. Sales trend analysis




**Sales trend analysis with Pages** :Pages functionality of Tableau is used to show how the sales pattern changes after every transaction.

**Product line analysis with Filters**:
Filters for different product lines are used to see unique trends for each category. Some categories may show a stronger seasonality effect or higher growth rates than others, providing insights into which product lines drive overall sales trends.

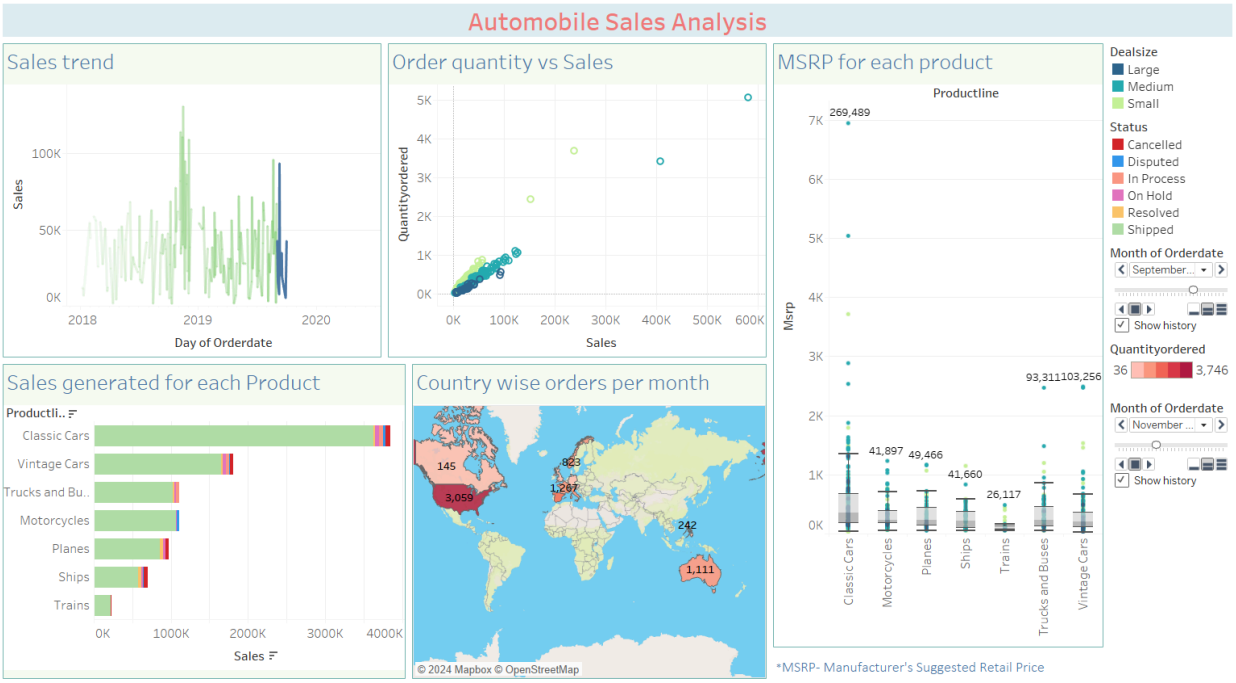
**The following are the observations from the above:**

- **Overall Upward Trend**: The visualization highlights a general increase in sales over time. This shows higher customer demand or expansion in product offerings. This could be influenced by various factors, such as market growth, improved customer acquisition, or effective marketing strategies.
- **Monthly Seasonality**: The sales pattern displays monthly seasonality, with recurring peaks and troughs each year. This indicates that sales vary depending on the time of year, possibly due to consumer buying behavior, product demand cycles, or seasonal events.
- **Peak Sales in Q4**: October to December show the highest sales volume annually. This corresponds with the increase in demand as consumers may tend to make more purchases in preparation for the holiday season.
- **Increased Sales in Early 2020**: The early months of 2020 showed an increase in sales compared to previous years. This may be due to strategic marketing, new product

launches, or changes in consumer behavior. Apart from this, the start of the COVID-19 pandemic in 2020 changed consumer habits. For some industries, this resulted in increased online purchases and changing buying patterns, which might have contributed to higher sales activity during this period.

# DASHBOARD

## Automobile-Parts Sales Analysis Dashboard




- Pricing and Order Volume: The 'Order Quantity vs Sales' chart shows the relationship between price (MSRP) and quantity ordered, which can help the automotive industry adjust their pricing to optimize sales.
- Country-wise Sales: The 'Country Wise Orders per Month' map visualizes sales by region, enabling targeted advertising and marketing efforts and region-specific pricing based on demand.
- Sales by Product Line: The 'Sales Generated for Each Product' chart highlights the high demand product lines, which will guide inventory management and marketing on high-demand products
- Sales trends are seen to fluctuate over time, suggesting possible seasonality or external factors influencing demand.
- A positive correlation is observed between order quantity and sales. This indicates that larger orders are likely driving higher revenue.

- Classic Cars generate the highest sales, followed by Vintage Cars and Trucks and Buses. This can confirm that these product lines contribute the most to revenue, probably due to higher demand or premium pricing. Trains generate the least sales, indicating either lower demand or limited product range in this category.
- There is a large variation in MSRP across different product lines. The higher MSRP for Classic Cars might be a contributing factor to their high sales revenue, as seen in the Sales Generated for Each Product chart.
- The United States has the highest number of orders i.e 3,056, followed by Australia-1,111, and France-1,267. This suggests that North America and many parts of Europe are major markets.

In summary, the dashboard investigates pricing, regional sales, and product performance to make data- driven strategic decisions in pricing, marketing, and inventory management.

## CONCLUSION

By analyzing pricing and order volume, regional sales patterns, and sales by product line, the automotive business can better understand key drivers of revenue. Targeted pricing strategies, targeted regional marketing efforts, and management of inventory on high-demand products will help optimize sales and profitability. This comprehensive approach can guide businesses to make decisions to offer products with market demand, resulting in a more data-driven business strategy.